

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2000年12月21日 (21.12.2000)

PCT

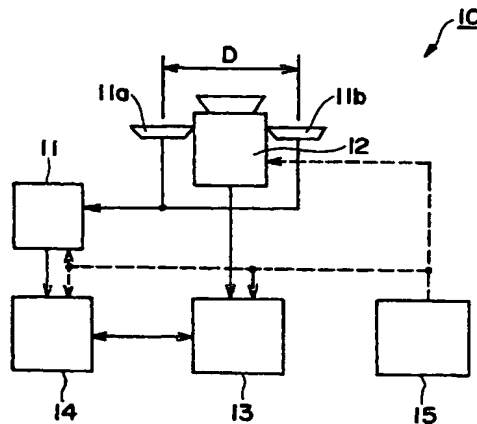
(10) 国際公開番号
WO 00/77537 A1

- (51) 国際特許分類: G01S 3/808, H04N 7/15, 7/18, H04R 1/40, 5/027
- (21) 国際出願番号: PCT/JP00/03695
- (22) 国際出願日: 2000年6月7日 (07.06.2000)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ: 特願平11/165182 1999年6月11日 (11.06.1999) JP
- (71) 出願人 (米国を除く全ての指定国について): 科学技術振興事業団 (JAPAN SCIENCE AND TECHNOLOGY CORPORATION) [JP/JP]; 〒332-0012 埼玉県川口市本町四丁目1番8号 Saitama (JP).
- (72) 発明者; および (75) 発明者/出願人 (米国についてのみ): 奥乃 博 (OKUNO, Hiroshi) [JP/JP]; 〒151-0066 東京都渋谷区西原2-10-9 Tokyo (JP). 北野宏明 (KITANO, Hiroaki) [JP/JP]; 〒350-0035 埼玉県川越市西小仙波町2-18-3 Saitama (JP). 中川友紀子 (NAKAGAWA, Yukiko) [JP/JP]; 〒211-0955 神奈川県川崎市幸区南加瀬5-1-17 Kanagawa (JP).
- (74) 代理人: 平山一幸 (HIRAYAMA, Kazuyuki); 〒160-0022 東京都新宿区新宿2-3-10 新宿御苑ビル6階 Tokyo (JP).
- (81) 指定国 (国内): US.
- (84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- 添付公開書類:
— 国際調査報告書

[続葉有]

(54) Title: METHOD AND APPARATUS FOR DETERMINING SOUND SOURCE

(54) 発明の名称: 音源同定装置及びその同定方法



(57) Abstract: A device for determining sound sources determines the positions of sound sources with high accuracy using sound information and image information. The position information is used to separate sounds from mixed sounds with high accuracy. The device (10) comprises sound collecting means (11) for processing the sounds collected by two microphones (11a, 11b) arranged at an interval of D opposed to a plurality of sound sources (A, B, C); camera means (12) for recording the images of the sound sources (A, B, C) continuously; sensor means for detecting objects; image processing means (13) for selecting position information on objects, which can be sound sources, from the images captured by the imaging means (12) and/or the direction information on the objects detected by the sensor means; sound processing means (14) for determining the positions of sound sources based on the sound information obtained by the sound collecting means (11) and the position information selected by the image processing means (13); and control means (15) for controlling the sound collecting means (11), the imaging means (12), the sensor means, the image processing means (13) and the sound processing means (14).

[続葉有]

WO 00/77537 A1



2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約:

本発明は、音響情報と画像情報とを用いて音源物体の位置をより高精度に同定して、その位置情報を用いて混合音から高精度で各音を分離し得るようにした、音源同定装置およびその同定方法を提供するもので、複数の音源（A、B、C）に対して所定間隔Dに配置した二つの集音マイク（11a）、（11b）で音を取り込んで処理する集音手段（11）と、上記複数の音源（A、B、C）を連続撮像する撮像手段（12）と、物体を検知する感知手段と、撮像手段（12）により撮像した画像及び感知手段により検知した物体の方向情報のいずれか、或いは両方から、音源となり得る物体に関する位置情報を選定する画像処理手段（13）と、集音手段（11）により集音した音響情報と画像処理手段（13）により選定した位置情報に基づいて、音源の位置を同定する音響処理手段（14）と、集音手段（11）、撮像手段（12）、感知手段、画像処理手段（13）及び音響処理手段（14）を制御する制御手段（15）とを含むように、音源同定装置（10）を構成する。

明 細 書

音源同定装置及びその同定方法

技術分野

本発明は、複数の音源からの画像情報と音響情報に基づいて、各音源を個別に同定するための音源同定装置及び方法に関するものである。

背景技術

従来から、混合音における音声、楽音といった特定の音について混合音から分離する方法の研究が行われている。例えば、入力音として音声を仮定する音声認識システムが知られている。また画像処理自体に関して、対象物の抽出に色、形状、動きを特徴として仮定するシステムが知られている。

しかしながら、音認識と画像処理を組み合わせた音認識システムはなく、また音声を仮定する音認識システムでは、マイクロフォンが口元にある時や他の音源が全く存在しない時にしか成立しない。

さらに複数の音源から調波構造を基に特定の音響信号を分離するとともに音源方向を求めるものもあるが、音源方向の精度は $\pm 10^\circ$ であり、隣接音源が近接している場合は音源の分離ができない。

また音源の数と同じ数の集音マイクを使用することにより、各集音マイクからの音響情報に基づいて、音源の同定を行なう方法も提案されている。これは音の強度と音源の位置を同定するものであるが、その周波数情報は方位軸に沿って拡散してしまい、良質な音源同定が困難である。さらにこのような方法では、音源の認識率を高めることは可能であるが、各音源がそれぞれ独立していることと、音源の数だけ集音マイクが必要であることから、コストが高くなる。

本発明は、このような従来の技術における欠点を解決するものであり、第1の目的として、音響情報と画像情報とを用いて音源物体の位置をより高精度に同定して、その位置情報を用いて混合音から高精度で各音を分離し得るようにした、音源同定装置を提供することにある。

さらに本発明の第2の目的として、音響情報と画像情報とを用いて音源物体の位置をより高精度に同定して、その位置情報を用いて混合音から高精度で各音を分離し得るようにする同定方法を提供することにある。

発明の開示

上記第1の目的を達成するために、本発明に係る音源同定装置によれば、複数の音源に対して所定間隔に配置した二つの集音マイクから成る集音手段と、上記複数の音源を連続撮像する撮像手段及び物体を検知する感知手段のいずれか或いは両方と、上記撮像手段により撮像された画像及び感知手段により検知した物体の方向情報のいずれか或いは両方から音源となり得る物体に関する位置情報を選定する画像処理手段と、集音手段により集音された音響情報と画像処理手段により選定された位置情報に基づいて、音源の位置を同定する音響処理手段と、上記集音手段、撮像手段、感知手段、画像処理手段及び音響処理手段を制御する制御手段とを具備することにより構成される。

さらに上記構成に加え、好ましくは、音響処理手段が特定時刻の音響情報だけを取り出す方向フィルタを備えている。

音響処理手段は、好ましくは、音源の概略位置情報を選定する機能を有している。

上記感知手段は、好ましくは、音源となり得る物体の磁気又は赤外線に基づいて検知する。

好ましくは、音源となり得る物体には、磁気を帯びた器材が設けられる。

このような構成により、本発明の音源同定装置では、集音マイクから得た音響情報により音源の位置を同定する際に、撮像手段によって撮像した画像情報と感知手段によって得た方向情報とに基づく位置情報を参考にして音源の方向を絞り込んでいる。

したがって本発明の音源同定装置では、動画像や物体の方向情報を用いて音源となりうる物体を特定するとともに、その位置情報と音響情報とを用いて音源分離を確実に行うことができる。

本発明の第2の目的を達成するために、本発明に係る音源同定方法によれば、

複数の音源に対して所定間隔に配置した二つの集音マイクから成る集音手段により集音する第一の段階と、第一の段階と同時に、上記複数の音源となり得る物体の連続撮像及び方向検知のいずれか或いは両方を行なう第二の段階と、第二の段階で撮像された画像及び方向検知のいずれか或いは両方から音源となり得る物体に関する位置情報を選定する第三の段階と、第一の段階により集音された音響情報及び第三の段階で選定された位置情報に基づいて、音源の位置を同定する第四の段階とを具備することにより構成される。

本発明の音源同定方法は、好ましくは、上記第一の段階により集音された音響情報のみに基づいて、音源の位置情報を概略的に選定する第五の段階を含んでおり、上記第三の段階が、この第五の段階により選定された概略位置情報に基づいて、前以て音源の方向を絞り込んで音源となり得る物体に関する位置情報を選定する。

本発明の音源同定方法は、好ましくは、上記第五の段階が、二つの集音マイクにより取得された音響情報の位相差及び強度差により、音源の方向を概略的に選定する。

本発明の音源同定方法は、好ましくは、上記第三の段階における音源となり得る物体に関する位置情報が、色及び形状のいずれか或いは両方を基準として選定される。

本発明の音源同定方法は、好ましくは、上記第四の段階が、第三の段階で選定された位置情報に基づいて、前以て設定された方向フィルタを選択することにより、各音源からの音響情報を取り出して、各音源の位置を同定する。

本発明の音源同定方法は、好ましくは、上記第四の段階または第五の段階が、第一の段階により得られた音響情報に基づいて、任意に分割した各周波数帯域の信号を基準として、音源の位置を選定する。

さらに音源となり得る物体に関する位置情報を、物体の動きを基準として選定するようにしてもよい。

また方向検知を磁気又は赤外線に基づいて検知することもできる。

これらの構成によれば、複数の音源に対して二つの集音マイクからなる集音手段によって音響情報を得ると共に、これらの音源を撮像手段によって撮像して画

像情報を得る。さらに音源の方向を磁気、赤外線に基づいて検知し、方向検知情報を得る。そして、音響処理手段が音響情報に基づいて、例えばその集音マイクにより取得された各音響情報の位相差及び強度差によって音源の位置を同定する際に、撮像手段によって得られた画像情報及び方向検知情報のいずれか、或いは両方に基づいて、例えばその色、形状、動き等により画像処理手段によって選定された音源となり得る物体に関する位置情報を参考にして音源の方向を絞り込んで、各周波数帯域の信号例えば調波構造を基準として音源の位置を同定するようにしている。したがって、音源の位置の同定に関して全方向に関する音響情報の処理が不要となり、より正確な音源の位置の同定を行なうことができると共に、処理情報量が少なく済み、処理時間を短縮することができる。

この場合、集音手段の二つの集音マイクによって、三つ以上の複数の音源の位置を同定することができるので、簡単な構成によって正確な音源の位置の同定を行なうことが可能になる。

また、上記第一の段階により集音された音響情報のみに基づいて、音源の位置情報を概略的に選定する第五の段階を含んでおり、第三の段階がこの第五の段階により選定された概略位置情報に基づいて、前以て音源の方向を絞り込んで音源となり得る物体に関する位置情報を選定するようになっている場合には、第三の段階による画像情報に基づく音源となり得る物体に関する位置情報の選定における処理情報量が軽減されるので、処理を簡単に行なうことができる。

上記第四の段階が、第三の段階で選定された位置情報に基づいて前以て設定された方向フィルタを選択することにより各音源からの音響情報を取り出して、各音源の位置を同定する場合には、ある方向に対する音源からの音響情報を取り出すための方向フィルタが前以て設定されているので、音源の位置の同定のための処理を円滑に行なうことができる。

図面の簡単な説明

本発明は以下の詳細な説明及び本発明の実施の形態を示す添付図面によって、よりよく理解されるものとなろう。なお、添付図面に示す実施例は本発明を特定又は限定することを意図するものではなく、単に説明及び理解を容易とするため

にだけ用いられるものである。

図中、

図 1 は本発明による音源同定装置の第 1 の実施例の構成を示す概略図である。

図 2 は図 1 の音源同定装置における撮像手段による撮像画面の一例を示す概略図である。

図 3 は図 1 の音源同定装置における撮像画面に対する説明図であり、(A) は音響処理手段による概略方向 A_0 、 B_0 、 C_0 を、また、(B) は画像処理手段による枠 A_1 、 B_1 、 C_1 を、さらに、(C) は画像処理手段による音源となり得る物体の位置情報 A_3 、 B_3 、 C_3 を、それぞれ示している。

図 4 は図 1 の音源同定装置における集音手段の二つの集音マイクと音源との距離差を示す説明図である。

図 5 は図 1 の音源同定装置における音響処理手段での方向フィルタの作用を示すグラフである。

図 6 は図 1 の音源同定装置における音響処理手段での同一音源からの二つの音響情報の取出しを示すグラフである。

図 7 は図 1 の音源同定装置における音響処理手段での方向フィルタによる各音源からの音響情報の取出しを示す説明図である。

図 8 は図 1 の音源同定装置における動作方法を示すフローチャートである。

図 9 は図 1 の音源同定装置における撮像手段による連続撮像画面の一部を示す図である。

図 10 は図 1 の音源同定装置における画像処理手段による種々の基準での音源となり得る物体の位置情報を示すグラフである。

発明を実施するための最良の形態

次に、本発明の音源同定装置及びその同定方法における最良の実施形態を図面を参照して詳細に説明する。

図 1 は本発明の音源同定装置の実施例を示している。

図 1 において、音源同定装置 10 は、集音手段 11 と、撮像手段 12 と、画像処理手段 13 と、音響処理手段 14 と、制御手段 15 と、を含んでいる。

上記集音手段 1 1 は、複数の音源（例えば三人の話者）に対して、所定間隔 D（図 1 参照）に配置した二つの集音マイク 1 1 a, 1 1 b で音を取り込んで処理している。これらの集音マイクの配置は適宜決定し得るが、図 1 に示した例では撮像手段 1 2 の両側、すなわち左右に設けられている。

上記撮像手段 1 2 は、例えば CCD（固体撮像素子）カメラから構成されており、図 2 に示すように、上記複数の音源（三人の話者 A, B, C）を含む画像を連続撮像するものである。

上記画像処理手段 1 3 は、撮像手段 1 2 により撮像された画像、例えば画像における色、形状又は動きに基づいて、音源となり得る物体に関する位置情報を選定するものである。なお、動きには振動なども含まれる。

この場合、画像処理手段 1 3 は、図 3（B）に示すように、撮像手段 1 2 により撮像された画像について、色（例えば人の肌の色）や高さ等に基づいて三人の話者 A, B, C についてそれぞれ枠 A 1, B 1, C 1 を設定し、図 3（C）に示すように、これらの枠 A 1, B 1, C 1 の中心位置 A 2, B 2, C 2（図 3 にてそれぞれ枠 A 1, B 1, C 1 内にて「+」図示）の水平座標 A 3, B 3, C 3 を音源となり得る物体に関する位置情報として選定する。

ここで、「音源となり得る物体」としたのは、必ずしも画像認識のみでは音源か否か不明だからである。

なお、画像処理手段 1 3 は、好ましくは画像処理の簡略化のために、上述の画像処理の前に、後述するように音響処理手段 1 4 で選定された各音源の概略方向 A 0, B 0, C 0（図 3（A）参照）が入力され、この概略方向 A 0, B 0, C 0 に絞り込んだ状態にて、即ちこれらの概略方向 A 0, B 0, C 0 の範囲内で上記画像処理を行なうことにより、音源となり得る物体に関する位置情報 A 3, B 3, C 3 を選定する。

上記音響処理手段 1 4 は、集音手段 1 1 のマイクにより集音された、例えば音響情報と画像処理手段 1 3 により選定された位置情報 A 3, B 3, C 3 に基づいて音源の位置を同定するものである。

音源の位置の同定は、音響情報に関しては左右の集音マイク 1 1 a, 1 1 b の音響情報の間の位相差及び強度差に基づいて行なわれる。

これは、図4に示すように、一つの音源からの音響情報は、その音源の方向 θ （ θ は、正面を0度とし、左にマイナス、右にプラスとする。）に対応して、左右の集音マイク11a, 11bから音源までの距離が d （ $d = D \cdot \sin \theta$ ）だけ異なることに起因して、集音マイク11a, 11bに達する音の位相が異なると共に、距離差 d による減衰により強度差が異なることを利用している。

なお、ここでは、音源の位置が不明であることから、音響処理手段14は、 $-90^\circ \leq \theta \leq +90^\circ$ の全角度範囲に亘って上記処理を行なう。この場合、処理作業の軽減のために、例えば θ に関して一定間隔、例えば5度間隔で処理を行なうようにしてもよい。

音響処理手段14は、先ず集音手段11からの左右の音響情報に基づいて、音源の概略方向 A_0 , B_0 , C_0 を選定する。これは、従来行なわれている音源同定と同じであって、 $\pm 10^\circ$ 程度の精度である。

そして、音響処理手段14は、この概略方向 A_0 , B_0 , C_0 を画像処理手段13に出力する。

さらに、音響処理手段14は、画像処理手段13から入力される位置情報 A_3 , B_3 , C_3 を参照して、これらの位置情報 A_3 , B_3 , C_3 の範囲内に絞り込んだ状態で、即ちこれらの位置情報 A_3 , B_3 , C_3 の近傍にて、再び音響情報に基づいて音源の位置を同定する。

この場合、音響処理手段14は、各音源A, B, Cに関してそれぞれ適宜の所謂方向フィルタを選択することにより音源の位置を同定する。

ここで、方向フィルタは、図5に示すように、特定時刻 t_0 の音響情報のみを取り出すために作成され、音源の方向に対する対照表として制御手段15内の補助記憶手段（図示せず）内に記憶されており、画像処理手段13からの位置情報 A_3 , B_3 , C_3 に基づいて、音響処理手段14が適宜の方向フィルタを選択して、上記補助記憶手段から読み出す。

これにより、図6に示すように一つの方向 θ に関して、ある時刻 t_1 における右側の音響情報に対して、位相差による遅延時間 Δt 後の時刻 t_2 （ $t_2 = t_1 + \Delta t$ ）における左側の音響情報を取り出すことにより、音源から同時に発して各集音マイク11a, 11bにより集音された音響情報を取得する。なお、 Δt

は負の場合もあり得る。

このようにして、ある程度正確な方向情報を持った各音源A, B, Cについて、音響処理手段14が方向フィルタを選択することにより、図7に示すように、混合音からそれぞれの音響情報を得ることができる。

なお、ここでは、音響処理手段14は、位置情報A3, B3, C3により音源の方向がある程度絞り込まれているので、 θ の全角度範囲($-90^\circ \leq \theta \leq +90^\circ$)について処理を行なう必要はなく、位置情報A3, B3, C3に関して所定の角度範囲で処理を行なえばよい。

上記制御手段15は、例えばコンピュータ等から構成されており、上記集音手段11, 撮像手段12, 画像処理手段13及び音響処理手段14を制御すると共に、上述したように前以て設定された方向フィルタが補助記憶手段(図示せず)に記憶されている。

本発明実施形態による音源同定装置10は以上のように構成されており、図8に示すフローチャートに従って、以下に説明するように動作する。

即ち、図8において、先ずステップST1にて、制御手段15が集音手段11を制御することにより、集音手段11の各集音マイク11a, 11bで、音源A, B, Cからの音を集音すると同時に、ステップST2にて、制御手段15が撮像手段12を制御することにより、音源の画像を連続的に撮像する。

次に、ステップST3にて、制御手段15が音響処理手段14を制御することにより、音響処理手段14が、集音手段11で得た同じ音源の同じ二つの音の位相差及び強度差の音響情報に基づいて、音源の概略方向A0, B0, C0(図3(A)参照)を選定する。このとき位相差のある全ての調波構造を調べ、大体の音源分離を行っている。なお、任意に分割した各周波数帯域の信号の一例として調波構造を基準とした。

続いて、ステップST4にて、制御手段15が画像処理手段13を制御することにより、画像処理手段13が、撮像手段12から撮像画面に基づいて音響処理手段14からの概略方向の範囲内で、画像の色、形状等により音源となり得る物体に関する位置情報A3, B3, C3(図3(C)参照)を選定する。

その後、ステップST5にて、制御手段15が音響処理手段14を制御するこ

とにより、音響処理手段 14 が、画像処理手段 14 からの位置情報 A 3, B 3, C 3 に関する所定角度範囲内の集音手段による音響情報に基づいて、音源 A, B, C の位置を同定する。

最後にステップ S T 6 にて、音響処理手段 14 が方向フィルタを選択し、同じ音源の同じ音の特定の時間遅れを含んだ音響情報だけを取り出す。このとき誤った他の調波構造の音響情報は処理しないので誤差が減り、音源分離率が上がる。

このようにして、本発明実施形態による音源同定装置 10 によれば、音響処理手段 14 が、集音手段 11 からの音響情報だけでなく、撮像手段 12 により撮像された画像に基づいて、画像処理手段 13 により音源となり得る物体の位置情報 A 3, B 3, C 3 を参照しながら、音源の位置を同定するので、従来の集音手段 11 からの音響情報だけの場合には ± 10 度前後の精度であったのに対して、本発明実施例による音源同定装置 10 によれば、より正確に音源の位置を同定することができる。

また予め大体の音源分離をした音源情報を画像情報から得られた位置情報を基に音源の位置精度を高めているので、近接する音源であっても音源同定が確実にできる。

具体的には、撮像手段 12 により、音源である三人の話者の連続撮像を行なった場合、例えば図 9 に示すような画像が得られる。なお、図 9 は連続撮像された画像のうち、7, 51, 78 及び 158 番目のフレームを示している。

ここで、各話者の正確な顔の位置は、図 10 (A) に示すようになっている。これにより各話者は、 θ が -30 度、 0 度及び $+20$ 度付近に位置していることが明らかである。

これに対して、画像処理手段 13 が色のみを基準として画像処理を行なって音源となり得る物体の位置を選定した場合、図 10 (B) のグラフに示すように、撮像画面中の種々の物体をも音源となり得る物体として誤認しているが、色及び高さを基準として画像処理を行なった場合には、図 10 (C) のグラフに示すように誤認が減少している。

また、画像処理手段 13 が、音響処理手段 14 からの概略方向 A 0, B 0, C 0 を参照して、色を基準として画像処理を行なった場合には、図 10 (D) のグ

ラフに示すようにより一層誤認が減少している。

さらに、画像処理手段 1 3 が、音響処理手段 1 4 からの概略方向 A 0, B 0, C 0 を参照して、色及び高さを基準として画像処理を行なった場合には、図 1 0 (E) のグラフに示すように、図 1 0 (A) に示した正確な顔の位置と比較して遜色のない、すなわち、かなり正確な音源の位置情報が選択され得ることが明白である。

なお、上述した実施例においては、画像処理手段 1 3 は、撮像した連続画像に基づいて、音源となり得る物体の枠 A 1, B 1, C 1 の中心位置 A 2, B 2, C 2 の水平座標 A 3, B 3, C 3 を音源となり得る物体に関する位置情報としているが、水平垂直座標を音源となり得る物体に関する位置情報としてもよい。

また、上述した実施例においては、画像処理手段 1 3 は、撮像した連続画像に基づいて、色や形状（例えば高さ）等を基準として音源となり得る物体の位置情報を選定するようになっている。

さらに、上述した実施例においては、画像処理手段 1 3 は音響処理手段 1 4 からの概略方向 A 0, B 0, C 0 を参考にして画像処理を行なっているが、これに限らず、撮像手段 1 2 からの画像情報のみにより、音源となり得る物体の位置情報を選定するようにしてもよい。

音源の方向を検知する場合、音源のそれぞれに磁気を帯びた器材のアクティブバッジなどを装着し、感知手段である磁気検知装置を用いて磁気を発する方向を選定してもよい。これを音響処理手段にフィードバックし、音響処理手段はこの磁気検知装置から得られた方向を用いて方向フィルターを作成し、音源を分離するようにしてもよい。

さらに音源が例えば人の場合、熱線を発しているので赤外線センサにより音源の方向を検知するようにしてもよい。

以上述べたように、この発明によれば、音響情報に基づいて音源を同定する際に、画像情報と方向検知情報とに基づいて、音源となり得る物体に関する位置情報を参考にして音源の方向を絞り込んでいるので、音源の同定に関して全方向に関する音響情報の処理が不要となり、より正確な音源の同定を行なうことができると共に、処理情報量が少なく済み、処理時間を短縮することができる。それ

故、この発明によれば、二本のマイクロフォンで高精度で複数の音源を同定し得るようにした極めて優れた音 同定装置及び方法が提供される。

なお本発明は例示的な実施例について説明したものであり、本発明の要旨及び範囲を逸脱することなく、実施例での種々の変更、省略、追加が可能である。したがって本発明は実施例に限定されるものではなく、特許請求の範囲に記載された要素によって規定される範囲及びその均等範囲を包含するものとして理解されなければならない。

産業上の利用可能性

以上のように、本発明の音源同定装置及びその同定方法は、音響情報と画像情報に基づき音源物体の位置をより高精度に同定し、その位置情報を用いて混合音から各音をより高精度に分離し得る音源同定装置及びその同定方法として極めて有用である。

請 求 の 範 囲

1. 複数の音源に対して所定間隔に配置した二つの集音マイクで音を取り込んで処理する集音手段と、

上記複数の音源となり得る物体を連続撮像する撮像手段及び物体を検知する感知手段のいずれか、或いは両方と、

上記撮像手段により撮像された画像及び上記感知手段により検知した物体の方向情報のいずれか、或いは両方から、音源となり得る物体に関する位置情報を選定する画像処理手段と、

集音手段により集音された音響情報と画像処理手段により選定された位置情報に基づいて、音源の位置を同定する音響処理手段と、

上記集音手段、撮像手段、感知手段、画像処理手段及び音響処理手段を制御する制御手段とを備える、音源同定装置。

2. 前記音響処理手段が、特定時刻の音響情報だけを取り出す方向フィルタを備えていることを特徴とする、請求の範囲第1項に記載の音源同定装置。

3. 前記音響処理手段が、前記音源となり得る物体の概略位置情報を選定する機能を有していることを特徴とする、請求の範囲第1項又は第2項に記載の音源同定装置。

4. 前記感知手段が、前記音源となり得る物体の磁気に基づいて検知することを特徴とする、請求の範囲第1項～第3項のいずれかに記載の音源同定装置。

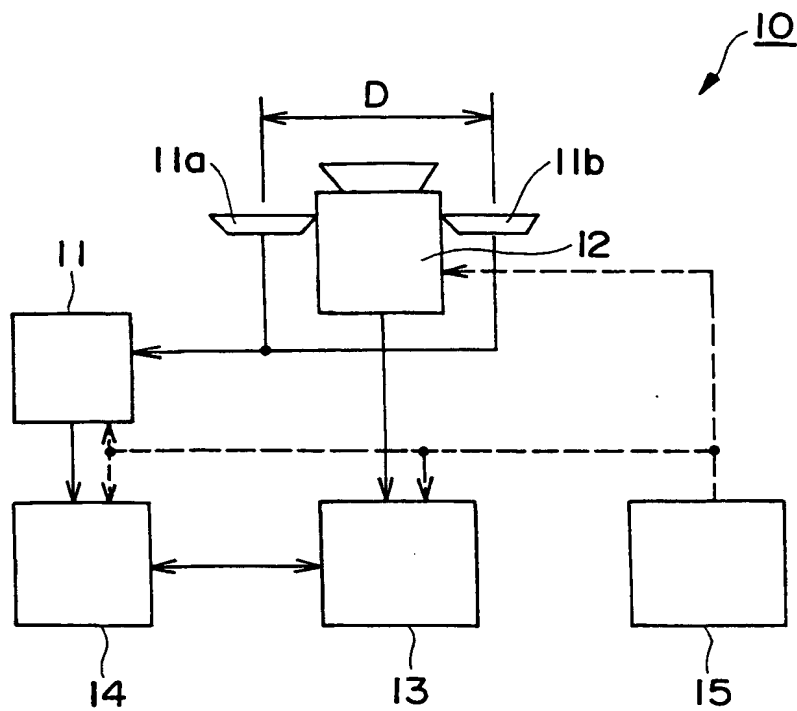
5. 前記感知手段が、前記音源となり得る物体の赤外線に基づいて検知することを特徴とする、請求の範囲第1項～第3項のいずれかに記載の音源同定装置。

6. 前記音源となり得る物体に磁気を帯びた器材を設けたことを特徴とする、請求の範囲第1項～第3項のいずれかに記載の音源同定装置。

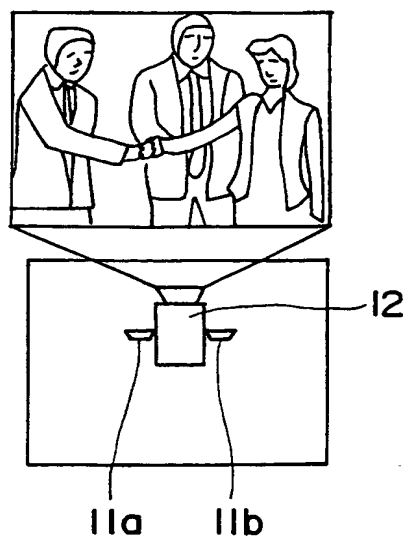
7. 複数の音源に対して所定間隔に配置した二つの集音マイクで音を取り込んで処理する集音手段により集音する第一の段階と、
第一の段階と同時に、上記複数の音源となり得る物体の連続撮像及び方向検知のいずれか、或いは両方を行なう第二の段階と、
第二の段階で撮像された画像及び方向検知のいずれか、或いは両方から、音源となり得る物体に関する位置情報を選定する第三の段階と、
第一の段階により集音された音響情報及び第三の段階で選定された位置情報に基づいて、音源の位置を同定する第四の段階と、
を含んでいることを特徴とする、音源同定方法。
8. 前記第一の段階により集音された音響情報のみに基づいて、音源の位置情報を概略的に選定する第五の段階を含んでおり、
前記第三の段階が、この第五の段階により選定された概略位置情報に基づいて、前以て音源の方向を絞り込んで音源となり得る物体に関する位置情報を選定することを特徴とする、請求の範囲第7項に記載の音源同定方法。
9. 前記第五の段階が、前記二つの集音マイクにより取得された音響情報の位相差及び強度差により、音源の方向を概略的に選定することを特徴とする、請求の範囲第8項に記載の音源同定方法。
10. 前記第三の段階における音源となり得る物体に関する位置情報が、色及び形状のいずれか、或いは両方を基準として選定されることを特徴とする、請求の範囲第7項～第9項の何れかに記載の音源同定方法。
11. 前記第四の段階が、前記第三の段階で選定された位置情報に基づいて、前以て設定された方向フィルタを選択することにより、各音源からの音響情報を取り出して、各音源の位置を同定することを特徴とする、請求の範囲第7項に記載の音源同定方法。

- 1 2. 前記第四の段階または第五の段階が、前記第一の段階により得られた音響情報に基づいて、任意に分割した各周波数帯域の信号を基準として、音源の位置を選定することを特徴とする、請求の範囲第 7 項～第 1 1 項の何れかに記載の音源同定方法。
- 1 3. 前記音源となり得る物体に関する位置情報を物体の動きを基準として選定していることを特徴とする、請求の範囲第 7 項～第 9 項, 第 1 1 項, 第 1 2 項のいずれかに記載の音源同定方法。
- 1 4. 前記方向検知を磁気に基づいて検知することを特徴とする、請求の範囲第 7 項～第 1 3 項のいずれかに記載の音源同定方法。
- 1 5. 前記方向検知を赤外線に基づいて検知することを特徴とする、請求の範囲第 7 項～第 1 3 項のいずれかに記載の音源同定方法。

第 1 図

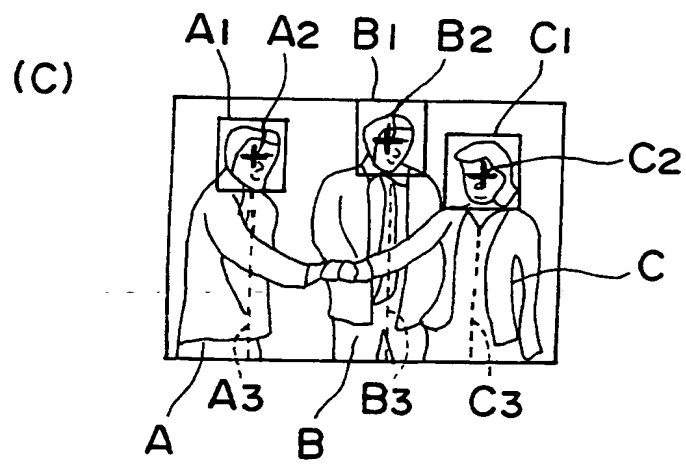
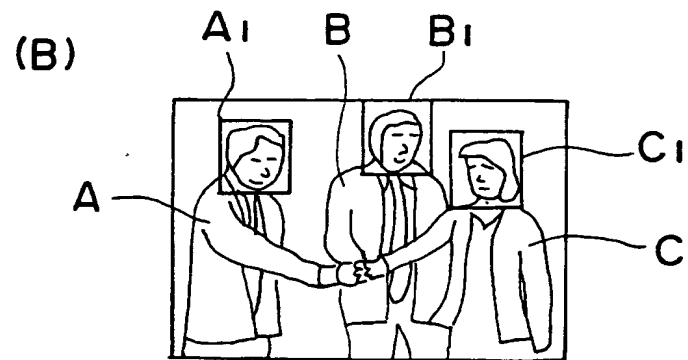
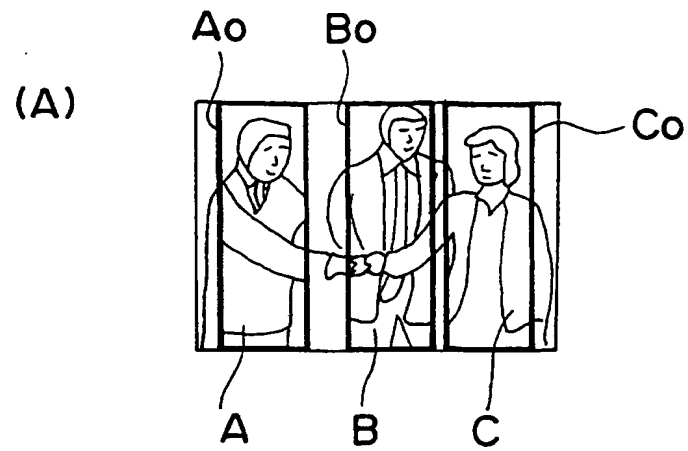


第 2 図



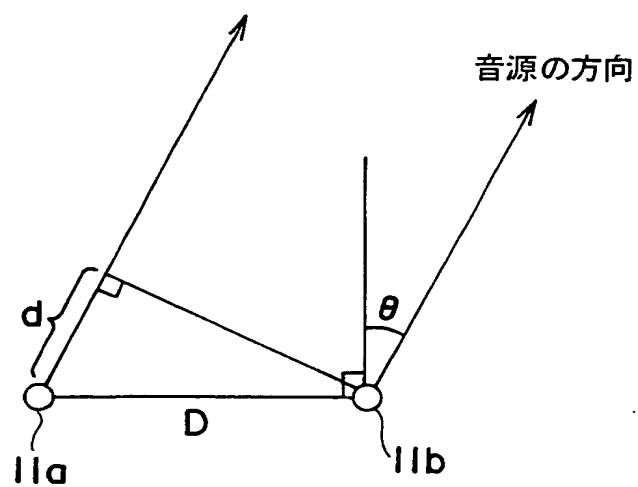
THIS PAGE BLANK (USPTO)

第 3 図

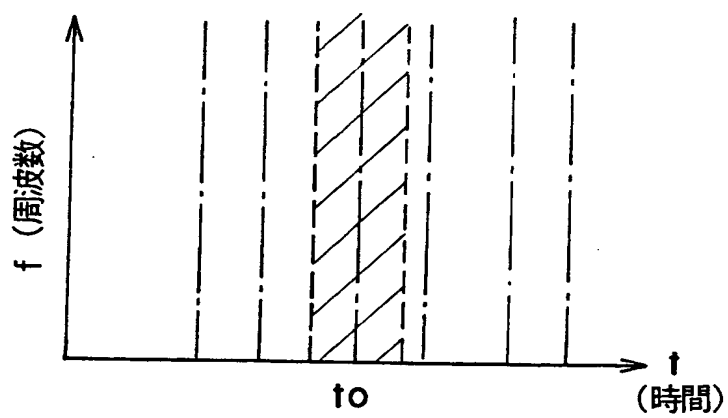


THIS PAGE BLANK (USPTO)

第 4 図

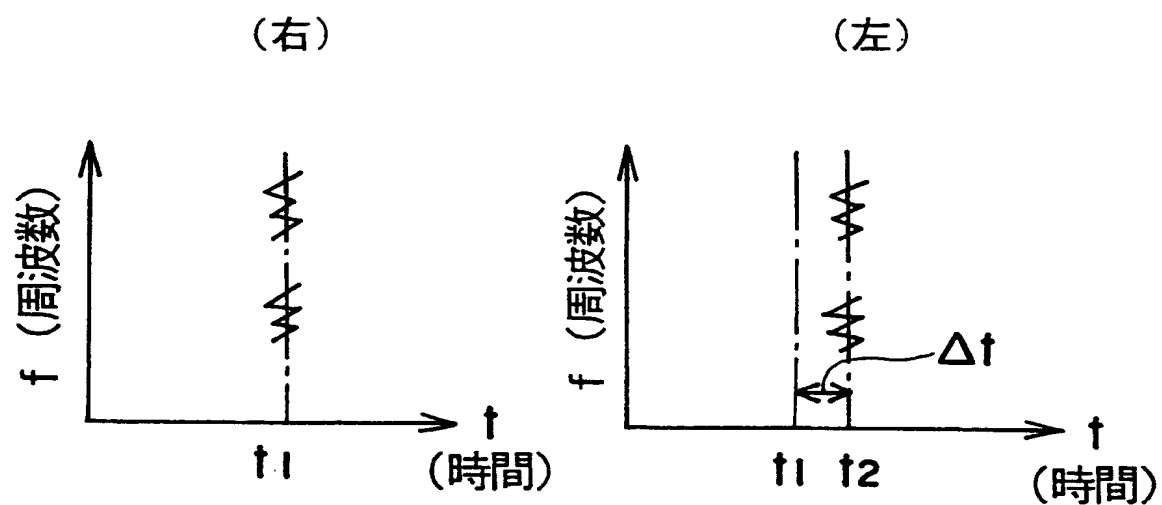


第 5 図



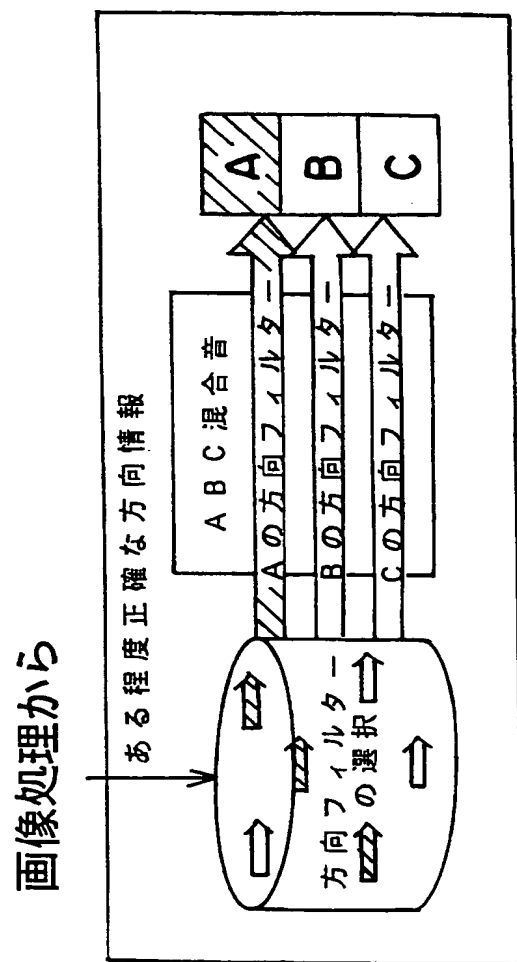
THIS PAGE BLANK (USPTO)

第 6 図



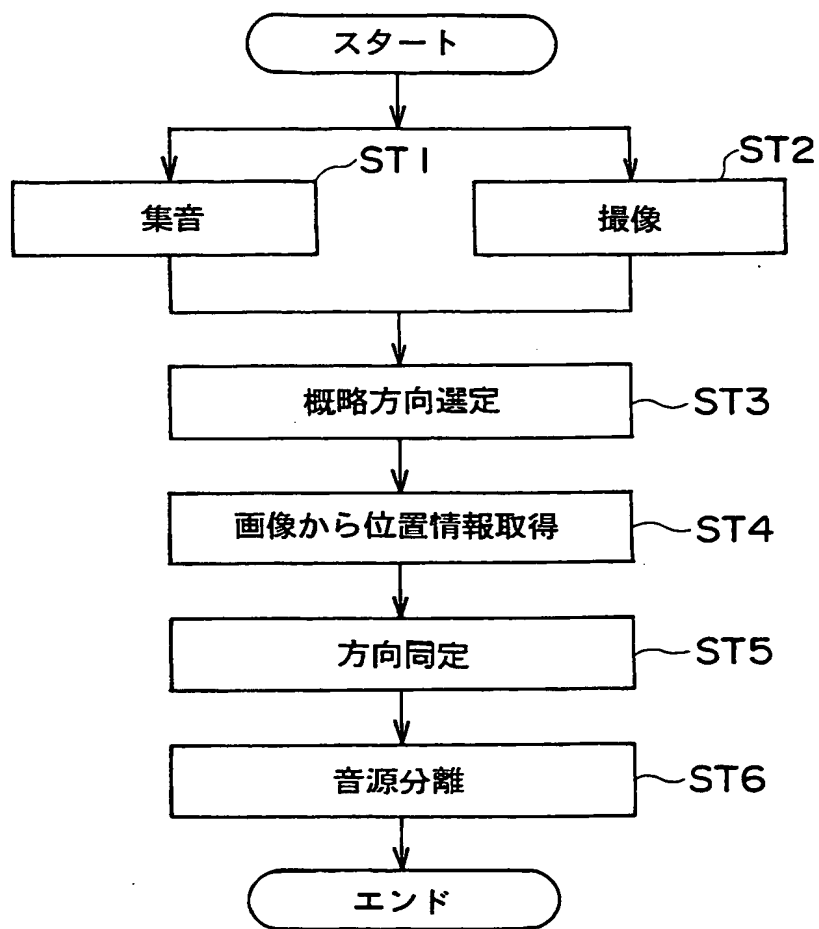
THIS PAGE BLANK (USPTO)

第 7 図



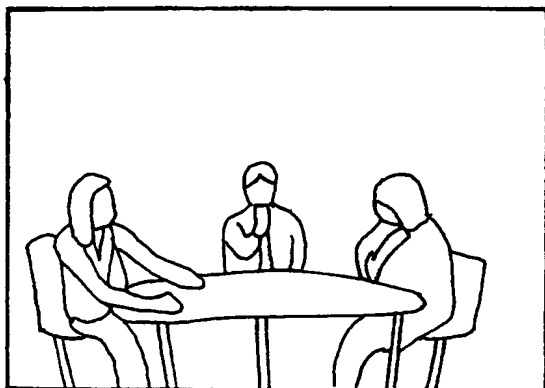
THIS PAGE BLANK (USPTO)

第 8 図

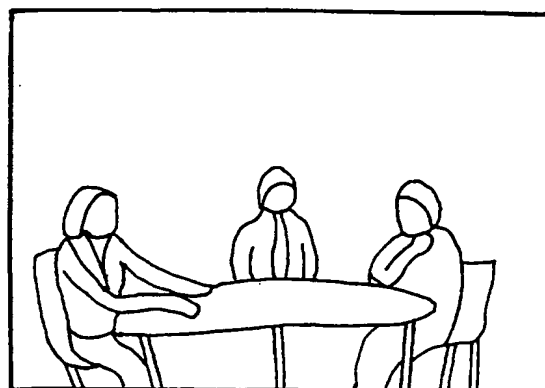


THIS PAGE BLANK (USPTO)

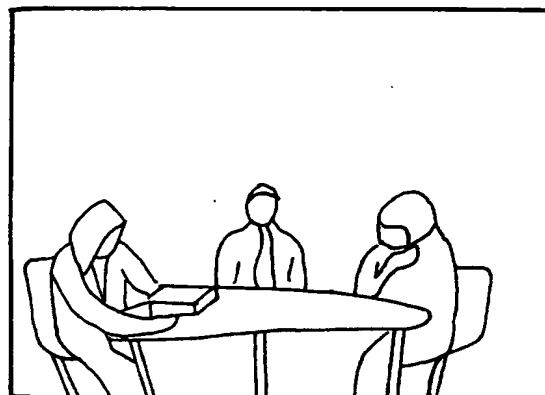
第 9 図



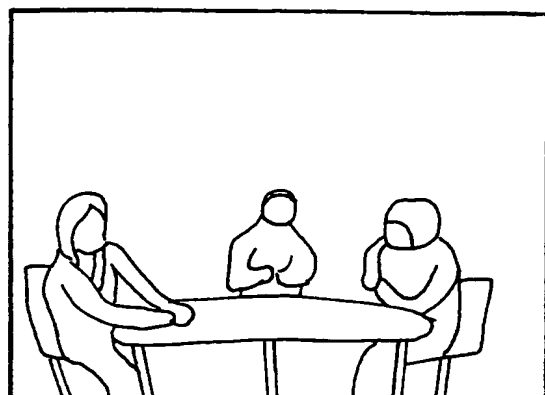
フレーム 7



フレーム 51



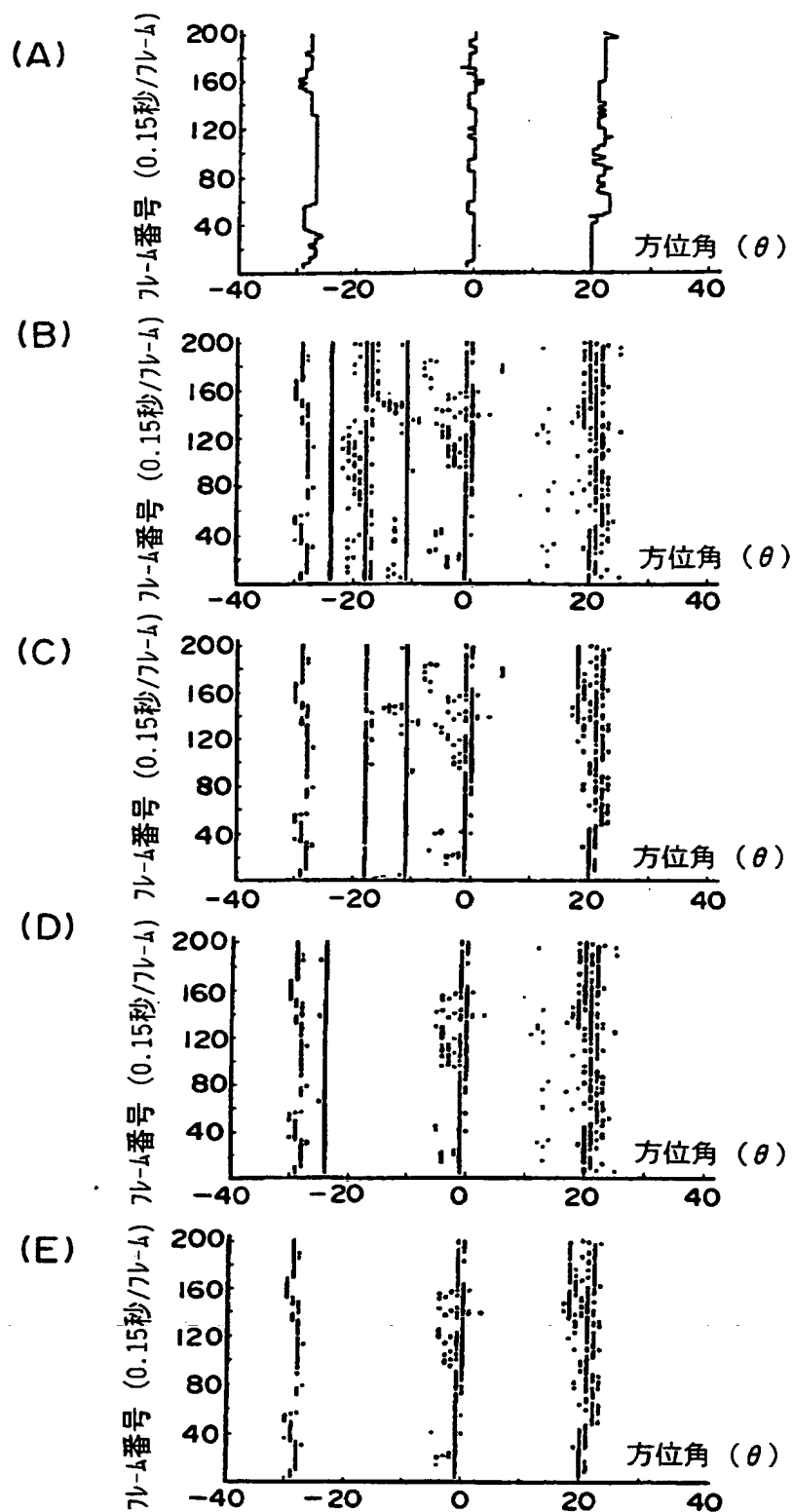
フレーム 78



フレーム 158

THIS PAGE BLANK (USPTO)

第 10 図



THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/03695

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G01S3/808
H04N7/15, H04N7/18, H04R1/40, H04R5/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G01S3/80-3/86, 5/18-5/30
H04N7/15, H04N7/18, H04R1/40, H04R5/027

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2000
Kokai Jitsuyo Shinan Koho 1971-2000 Jitsuyo Shinan Toroku Koho 1996-2000

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | JP, 11-41577, A (Fujitsu Limited), 12 February, 1999 (12.02.99), Full text; all drawings | 1, 3, 5, 7, 8, 152, |
| Y | Full text; all drawings (Family: none) | 4, 6, 9-14 |
| Y | JP, 9-33330, A (Nippon Telegr. & Teleph. Corp. <NTT>), 07 February, 1997 (07.02.97), Full text; all drawings (Family: none) | 2, 9, 12 |
| Y | JP, 8-251561, A (NEC Corporation), 27 September, 1996 (27.09.96), Par. No. [0044]; Figs. 1 to 8 & US, 5786846 | 4, 6, 14 |
| Y | JP, 8-286680, A (TAKENAKA CORPORATION), 01 November, 1996 (01.11.96), Full text; all drawings | 11 |
| Y | Par. No. [0010] (Family: none) | 10 |
| Y | JP, 5-244587, A (Mitsubishi Electric Corporation), 21 September, 1993 (21.09.93), Full text; all drawings (Family: none) | 10 |

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

| | |
|--|---|
| <p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> | <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p> |
|--|---|

Date of the actual completion of the international search
01 September, 2000 (01.09.00)

Date of mailing of the international search report
12 September, 2000 (12.09.00)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/03695

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| Y | JP, 10-313497, A (Nippon Telegr. & Teleph. Corp. <NTT>), 24 November, 1998 (24.11.98), Full text; all drawings & EP, 831458 | 2,11,12 |
| Y | JP, 10-51889, A (Toshiba Corporation), 20 February, 1998 (20.02.98), Full text; all drawings (Family: none) | 2,11,12 |
| Y | JP, 7-39000, A (Kazumoto SUZUKI), 07 February, 1995 (07.02.95), Full text; all drawings (Family: none) | 2,11,12 |
| A | JP, 6-105306, A (Funai Denki Kenkyusho K.K.), 15 April, 1994 (15.04.94), Full text; all drawings (Family: none) | 5,15 |

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G01S3/808

H04N7/15, H04N7/18, H04R1/40, H04R5/027

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G01S3/80-3/86, 5/18-5/30

H04N7/15, H04N7/18, H04R1/40, H04R5/027

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922 - 1996年

日本国公開実用新案公報 1971 - 2000年

日本国登録実用新案公報 1994 - 2000年

日本国実用新案登録公報 1996 - 2000年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

WPI

C. 関連すると認められる文献

| 引用文献の カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示 | 関連する 請求の範囲の番号 |
|-----------------|--|------------------------------------|
| X Y | J P, 11-41577, A (富士通株式会社) 12. 2月. 1999 (12. 02. 99) 全文, 全図 全文, 全図 (ファミリーなし) | 1, 3, 5, 7, 8, 15 2, 4, 6, 9-14 |
| Y | J P, 9-33330, A (日本電信電話株式会社) 7. 2月. 1997 (07. 02. 97) 全文, 全図 (ファミリーなし) | 2, 9, 12 |

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

01. 09. 00

国際調査報告の発送日

12.09.00

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

長浜 義憲

2 S

7719

電話番号 03-3581-1101 内線 3256

| C (続き) . 関連すると認められる文献 | | |
|-----------------------|--|------------------|
| 引用文献の カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示 | 関連する 請求の範囲の番号 |
| Y | J P, 8-251561, A (日本電気株式会社) 27. 9月. 1996 (27. 09. 96) 段落番号【0044】, 第1-8図 & US, 5786846 | 4, 6, 14 |
| Y | J P, 8-286680, A (株式会社竹中工務店) 1. 11月. 1996 (01. 11. 96) 全文, 全図 | 11 |
| Y | 段落番号【0010】 (ファミリーなし) | 10 |
| Y | J P, 5-244587, A (三菱電機株式会社) 21. 9月. 1993 (21. 09. 93) 全文, 全図 (ファミリーなし) | 10 |
| Y | J P, 10-313497, A (日本電信電話株式会社) 24. 11月. 1998 (24. 11. 98) 全文, 全図 & EP, 831458 | 2, 11, 12 |
| Y | J P, 10-51889, A (株式会社東芝) 20. 2月. 1998 (20. 02. 98) 全文, 全図 (ファミリーなし) | 2, 11, 12 |
| Y | J P, 7-39000, A (鈴木 一元) 7. 2月. 1995 (07. 02. 95) 全文, 全図 (ファミリーなし) | 2, 11, 12 |
| A | J P, 6-105306, A (株式会社船井電機研究所) 15. 4月. 1994 (15. 04. 94) 全文, 全図 (ファミリーなし) | 5, 15 |

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/03695

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G01S3/808

H04N7/15, H04N7/18, H04R1/40, H04R5/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G01S3/80-3/86, 5/18-5/30

H04N7/15, H04N7/18, H04R1/40, H04R5/027

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

| | | | |
|---------------------------|-----------|----------------------------|-----------|
| Jitsuyo Shinan Koho | 1922-1996 | Toroku Jitsuyo Shinan Koho | 1994-2000 |
| Kokai Jitsuyo Shinan Koho | 1971-2000 | Jitsuyo Shinan Toroku Koho | 1996-2000 |

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPI

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| X | JP, 11-41577, A (Fujitsu Limited), 12 February, 1999 (12.02.99), Full text; all drawings | 1, 3, 5, 7, 8, 152, |
| Y | Full text; all drawings (Family: none). | 4, 6, 9-14 |
| Y | JP, 9-33330, A (Nippon Telegr. & Teleph. Corp. <NTT>), 07 February, 1997 (07.02.97), Full text; all drawings (Family: none) | 2, 9, 12 |
| Y | JP, 8-251561, A (NEC Corporation), 27 September, 1996 (27.09.96), Par. No. [0044]; Figs. 1 to 8 & US, 5786846 | 4, 6, 14 |
| Y | JP, 8-286680, A (TAKENAKA CORPORATION), 01 November, 1996 (01.11.96), Full text; all drawings | 11 |
| Y | Par. No. [0010] (Family: none) | 10 |
| Y | JP, 5-244587, A (Mitsubishi Electric Corporation), 21 September, 1993 (21.09.93), Full text; all drawings (Family: none) | 10 |

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
01 September, 2000 (01.09.00)

Date of mailing of the international search report
12 September, 2000 (12.09.00)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/03695

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| Y | JP, 10-313497, A (Nippon Telegr. & Teleph. Corp. <NTT>), 24 November, 1998 (24.11.98), Full text; all drawings & EP, 831458 | 2,11,12 |
| Y | JP, 10-51889, A (Toshiba Corporation), 20 February, 1998 (20.02.98), Full text; all drawings (Family: none) | 2,11,12 |
| Y | JP, 7-39000, A (Kazumoto SUZUKI), 07 February, 1995 (07.02.95), Full text; all drawings (Family: none) | 2,11,12 |
| A | JP, 6-105306, A (Funai Denki Kenkyusho K.K.), 15 April, 1994 (15.04.94), Full text; all drawings (Family: none) | 5,15 |